

The error structure of the SMAP single and dual channel soil moisture retrievals

Jianzhi Dong¹, Wade Crow¹, Rajat Bindlish²

¹USDA Hydrology and Remote Sensing Laboratory, Beltsville, Maryland, USA

²NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Key Points:

- Retrieval errors for the SMAP single and dual polarization algorithms are compared.
- The single channel algorithm provides poorer estimates for regions with high inter-annual variability in vegetation opacity.
- The single channel algorithm has stronger auto-correlated errors, which is attributed to its use of a climatology for vegetation opacity.

Abstract

Knowledge of the temporal error structure for remotely sensed surface soil moisture retrievals can improve our ability to exploit them for hydrologic and climate studies. This study employs a triple collocation analysis to investigate both the total variance and temporal auto-correlation of errors in Soil Moisture Active and Passive (SMAP) products generated from two separate soil moisture retrieval algorithms, the vertically-polarized brightness temperature based Single Channel Algorithm (SCA-V, the current baseline SMAP algorithm) and the Dual Channel Algorithm (DCA). A key assumption made in SCA-V is that real-time vegetation opacity can be accurately captured using only a climatology for vegetation opacity. Results demonstrate that, while SCA-V generally outperforms DCA, SCA-V can produce larger total errors when this assumption is significantly violated by inter-annual variability in vegetation health and biomass. Furthermore, larger auto-correlated errors in SCA-V retrievals are found in areas with relatively large vegetation opacity deviations from climatological expectations. This implies that a significant portion of the auto-correlated error in SCA-V is attributable to the violation of its vegetation opacity climatology assumption and suggests that utilizing a real (as opposed to climatological) vegetation opacity time series in the SCA-V algorithm would reduce the magnitude of auto-correlated soil moisture retrieval errors.

1 Introduction

The Soil Moisture Active and Passive (SMAP) mission [Entekhabi *et al.*, 2010, 2014] has provided global surface soil moisture using an L-band (1.413 GHz) radiometer since March 31, 2015. Compared with C- and X-band microwave remote sensing, the L-band microwave emission has increased sensitivity to soil moisture and improved vegetation penetration [Entekhabi *et al.*, 2014]. Hence, SMAP retrieved soil moisture products are expected to significantly benefit hydrologic, climate, and land surface/atmosphere coupling studies [Entekhabi *et al.*, 2010; Brown *et al.*, 2013; Koster *et al.*, 2016].

There are two primary SMAP soil moisture retrieval algorithms: 1) the single channel algorithm (SCA) and 2) the dual channel algorithm (DCA). The SCA version, which uses vertically-polarized brightness (SCA-V), is the current SMAP baseline algorithm. Both SCA-V and DCA use the same radiative transfer model but a key difference is their treatment of vegetation opacity (τ). In SCA-V, τ is assumed to be proportional to vegetation water content (VWC), which, in turn, is acquired via the application of an empirical

function to satellite-derived normalized vegetation difference index (NDVI) values. However, the real-time application of this approach is complicated by the latency for Moderate Resolution Imaging Spectroradiometer (MODIS) NDVI products due to the compositing of daily NDVI retrievals to account for cloud cover. In addition, real-time NDVI estimates typically contain substantial random noise that can be transferred into soil moisture retrievals. Therefore, in order to maintain the 24-hour latency goal of SMAP Level 2 soil moisture products and suppress random errors, the real-time application of SCA-V within the SMAP data product stream uses a (seasonally-varying) climatological value of VWC calculated from a 2000 to 2010 baseline [Chan *et al.*, 2013]. Soil moisture is then retrieved using vertically-polarized brightness temperature and a value of τ obtained from this VWC climatology [O'Neill *et al.*, 2015].

In contrast, DCA feeds an initial guess for τ and soil moisture into a forward radiative transfer model. These initial guesses are then iteratively adjusted to minimize the difference between simulated and observed dual-polarized (i.e., vertical and horizontal) brightness temperature [O'Neill *et al.*, 2015]. Consequently, DCA treats both τ and soil moisture as free parameters to be iteratively solved for in a simultaneous manner.

Due to the low-frequency nature of variations in vegetation health, VWC inter-annual anomalies (i.e., the difference between the VWC seasonal climatology and real-time VWC) are highly auto-correlated in time. For SCA-V, this potentially leads directly to auto-correlated τ errors (since VWC anomalies are neglected) and, by extension, auto-correlated errors in SCA-V surface soil moisture retrievals. Therefore, while numerical tests and field validations results suggest that SCA generally outperforms DCA [Entekhabi *et al.*, 2014; Chan *et al.*, 2016, 2017], the real-time SCA-V algorithm may be prone to auto-correlated error in the presence of large VWC inter-annual variability.

The presence of auto-correlated error in soil moisture retrievals has a number of important implications. For instance, the analysis of soil moisture memory [e.g. Koster and Suarez, 2001; McColl *et al.*, 2017] should be conducted on a time scale that is longer than the soil moisture error auto-correlation length to avoid biased estimates. Likewise, deriving root-zone soil moisture estimates via the processing of surface soil moisture retrievals through an exponential filter [Albergel *et al.*, 2008] is more vulnerable to auto-correlated errors than temporally white errors. This is because the exponential filter is essentially a weighted average of surface soil moisture observations, and while white noise can be ef-

fectively reduced by averaging, auto-correlated errors will persist after temporal averaging and therefore play a dominant role in root-zone soil moisture estimation errors. Finally, information concerning soil moisture error auto-correlation is useful for the assimilation of surface soil moisture retrievals into land surface models [Crow and Van den Berg, 2010]. In an optimal data assimilation approach, auto-correlated observation error should be accounted for via an augmented state vector [Crow and Yilmaz, 2014]. However, existing attempts to characterize SMAP soil moisture retrieval errors [e.g. Chan et al., 2016; Colliander et al., 2017; Montzka et al., 2017; Ray et al., 2017] have not directly addressed the degree to which these errors are auto-correlated in time.

Therefore, the primary goal of this study is to investigate both the total variance and temporal autocorrelation coefficient of errors in SMAP Level 2 soil moisture retrievals generated using both the SCA-V and DCA retrieval algorithms. The analysis will be conducted using an existing triple collocation (TC) approach [Scipal et al., 2008; Zwieback et al., 2013]. In doing so, we seek to quantify both the presence of auto-correlated soil moisture errors and attribute the spatial patterns of these errors to specific aspects of both retrieval algorithms.

2 Methods and data collection

2.1 TC based error analysis

Suppose three independent soil moisture products are available, and their zero mean anomalies (\mathbf{x} , \mathbf{y} and \mathbf{z}) are linearly related to the true soil moisture anomaly (θ) as [Gruber et al., 2016]:

$$\mathbf{x} = \alpha_x \theta + \epsilon_x \quad (1)$$

$$\mathbf{y} = \alpha_y \theta + \epsilon_y \quad (2)$$

$$\mathbf{z} = \alpha_z \theta + \epsilon_z \quad (3)$$

where the α_x , α_y and α_z represent the scale differences between the products and the truth, and the ϵ_x , ϵ_y and ϵ_z represent the errors of each product. Note that the soil moisture anomaly in this study is defined as an anomaly relative to a long-term climatology [Chen et al., 2017], which is calculated for each-day-of-year (DOY) by averaging all available soil moisture estimates within a 31-day window centered on each DOY across multiple years. Anomalies are then calculated as the differences between the actual daily soil moisture time series and this long-term climatology. Using a shorter window length may

lead to insufficient sample sizes in calculating the climatology. On the contrary, an extremely long window length results in an over-smoothed seasonal climatology. A 31-day window length is used here as a trade-off between these two considerations. Results presented in the supporting materials confirm that key results presented here are relatively insensitive to variations in this window length.

In order to eliminate scale differences, TC analysis typically takes one product as a reference and scales the other two products to this reference. Here, we take \mathbf{x} as a reference product for illustration. Assuming the errors of the three products are mutually independent and orthogonal to the truth, the scale parameters can then be calculated as:

$$\frac{\alpha_x}{\alpha_y} = \frac{\mathbf{x}^T \mathbf{z}}{\mathbf{y}^T \mathbf{z}} \quad (4)$$

$$\frac{\alpha_x}{\alpha_z} = \frac{\mathbf{x}^T \mathbf{y}}{\mathbf{z}^T \mathbf{y}} \quad (5)$$

Consequently, the scaled product $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ can be expressed as:

$$\mathbf{x}^* = \mathbf{x} \quad (6)$$

$$\mathbf{y}^* = \frac{\alpha_x}{\alpha_y} \mathbf{y} \quad (7)$$

$$\mathbf{z}^* = \frac{\alpha_x}{\alpha_z} \mathbf{z} \quad (8)$$

The variance of the errors (σ_x^2 , σ_y^2 and σ_z^2) in the scaled products can be estimated as:

$$\sigma_x^2 = \langle (\mathbf{x}^* - \mathbf{y}^*)^T (\mathbf{x}^* - \mathbf{z}^*) \rangle \quad (9)$$

$$\sigma_y^2 = \langle (\mathbf{y}^* - \mathbf{x}^*)^T (\mathbf{y}^* - \mathbf{z}^*) \rangle \quad (10)$$

$$\sigma_z^2 = \langle (\mathbf{z}^* - \mathbf{x}^*)^T (\mathbf{z}^* - \mathbf{y}^*) \rangle \quad (11)$$

where $\langle \cdot \rangle$ denotes temporal averaging. Likewise, the serial lag-1 temporal autocovariance of the observation sequences (L_x , L_y and L_z) can be estimated as [Zwieback *et al.*, 2013]:

$$L_x = \langle (\mathbf{x}^* - \mathbf{y}^*)^T (\mathbf{x}_L^* - \mathbf{z}_L^*) \rangle \quad (12)$$

$$L_y = \langle (\mathbf{y}^* - \mathbf{x}^*)^T (\mathbf{y}_L^* - \mathbf{z}_L^*) \rangle \quad (13)$$

$$L_z = \langle (\mathbf{z}^* - \mathbf{x}^*)^T (\mathbf{z}_L^* - \mathbf{y}_L^*) \rangle \quad (14)$$

where the subscript “L” denotes a serial lagged-1 vector. The ratio of the error autocovariance (i.e. L_x to L_z) and the error variance (i.e. σ_x^2 to σ_z^2) is the correlation coefficient of the observation errors.

As shown in *Dorigo et al.* [2010], the TC analysis is less robust for regions with low signal-to-noise ratio (SNR, the ratio of the variance of the soil moisture signal to the variance of the total soil moisture retrieval error [*Gruber et al.*, 2016]). In particular, an examination of equations (4) and (5) reveals that the TC scaling factors are particularly vulnerable to sampling errors when the SNR of individual soil moisture products is small. Therefore, in order to provide a robust analysis, this study considered TC results only for grids where the signal to noise ratio (SNR) of the remotely sensed soil moisture retrievals are significantly greater than 0.1 [-] (at $p = 0.05$ confidence). Following *Draper et al.* [2013], a 1500-member bootstrapping analysis was used to estimate the empirical sampling distribution of SNR and, by extension, the p value of $\text{SNR} > 0.1$ [-] for each grid. Each bootstrapping member was constructed by resampling the original dataset with replacement to preserve its sample size.

It should be stressed that the irregular temporal sampling of soil moisture remote-sensing data complicates the definition of a specific time scale time for derived lag-1 temporal auto-covariances. Although auto-covariance estimates for a specific temporal lag scale can be sampled using daily pairs with certain (fixed) temporal lags [*Zwieback et al.*, 2013], this requires discarding large amounts of data and is therefore infeasible for the (relatively short) SMAP data record. Instead, we utilized all available serial data pairs (regardless of their time lag) in calculating a serial lag-1 error auto-covariance using equations (12 - 14). For the remote-sensing data utilized here (see below), the mean time lag of these serial pairs (i.e., that mean time scale associated with a “lag-1” serial auto-covariance analysis) is 4.8 [days] with a standard deviation of 0.5 [days]. However, as noted above, this study primarily focuses on the differences between SCA-V and DCA error structures. Since the two retrieval products have exactly the same temporal sampling, time scale ambiguities associated with their irregular temporal sampling will not impact these differences.

2.2 SMAP L2 passive soil moisture data

The Level 2 SMAP soil moisture product (SPL2SMP) provides soil moisture retrievals using both SCA-V and DCA within a 36-km EASE2 grid. As noted earlier, the key difference between SCA-V and DCA lies in their treatment of vegetation opacity (τ). In SCA-V, a NDVI climatology is first calculated using data collected from 2000 to 2010.

Then, vegetation water content (VWC) is calculated using the NDVI climatology as:

$$\text{VWC} = \left(1.9134\text{NDVI}^2 - 0.3215\text{NDVI} \right) + W_{\text{stem}} \frac{\text{NDVI}_{\text{max}} - \text{NDVI}_{\text{min}}}{1 - \text{NDVI}_{\text{min}}} \quad (15)$$

where W_{stem} is a stem factor estimated using MODIS International Geosphere-Biosphere Programme (IGBP) land cover types [O'Neill *et al.*, 2015], and NDVI_{max} and NDVI_{min} are the annual maximum and minimum NDVI at a given location. The vegetation opacity (τ) is then calculated as:

$$\tau = b \times \text{VWC} \quad (16)$$

where b is a temporally constant proportionality value. In SMAP, the b value is determined using a look-up table based on IGBP land cover types [O'Neill *et al.*, 2015]. In DCA, soil moisture and VWC are simultaneously estimated by minimizing the differences between simulated and observed vertically and horizontally polarized brightness temperature [O'Neill *et al.*, 2015]. Here, SCA-V and DCA retrievals from the 6 AM local solar time descending pass during the period between March 31, 2015 and April 30, 2017 (with a revisit time of 2 to 3 days) were used. Full details about both algorithms can be found in O'Neill *et al.* [2015]. As noted above, SCA-V is the current baseline algorithm for all official SMAP Level 2 and 3 radiometer-based soil moisture products.

2.3 Advanced Scatterometers (ASCAT) soil moisture data

The Advanced Scatterometer (ASCAT) sensor onboard the Meteorological Operational-B (MetOp-B) satellite measures C-band (5.3 GHz) radar backscatter. The ASCAT Level 2 (v5) soil moisture index product used here has a spatial resolution of 25-km and is retrieved using a change-detection algorithm developed by the Vienna University of Technology [Wagner *et al.*, 1999; Naeimi *et al.*, 2009]. A vegetation climatology is used for removing vegetation impacts during the near real-time retrieval. Here, the ASCAT soil moisture retrievals were resampled onto a 36-km EASE2 grid consistent with SPL2SMP. To minimize the temporal differences between SMAP and ASCAT retrievals, only descending ASCAT data (9:30 AM local solar time overpass) were used in this study.

2.4 NLDAS 2-Noah data

A third independent soil moisture product was collected from the Noah model run quasi-operationally as part of Phase 2 of the North American Land Data assimilation System (NLDAS-2) experiment [Xia *et al.*, 2012]. Soil moisture estimates for the top layer

of the NOAH model (10 cm in depth) at 6 AM were utilized directly for the TC analysis. This hourly soil moisture product has a spatial resolution of 0.25 degrees, with units of kg/m^2 . To match the SMAP spatial grid, we resampled this Noah soil moisture product onto a 36-km EASE2 grid and converted the units of the resampled product from kg/m^2 into volumetric soil water content (m^3/m^3). The NOAH modeled soil moisture was selected as the reference soil moisture product in the TC error analysis.

2.5 NDVI data

As noted above, climatological values of NDVI are used for calculating τ in the SCA-V soil moisture retrieval algorithm. Therefore, the contribution of τ error to the (total) error presented in SCA-V retrievals is due (in part) to the deviation of τ from its climatological expectation. To estimate this quantity, MODIS 16-day NDVI composite (MOD13C1) data with a spatial resolution of 0.05-degrees were retrieved from the on-line NASA Land Processes Distributed Active Archive Center (LP DAAC) for the period of 31 March, 2015 to April 30, 2017. These NDVI data were then resampled onto a 36-km EASE2 grid. Following *Chan et al.* [2013], days without NDVI data were filled via piece-wise linear interpolation. To be consistent with soil moisture anomalies, the same approach (see section 2.1) was applied to calculate the NDVI climatology for each DOY. Using equation (15), this NDVI climatology was then converted into a VWC climatology and, subsequently, into a τ climatology using equation (16). In parallel, benchmark values for τ were calculated directly using the interpolated NDVI time series. The difference between climatological and benchmark τ values represent an inter-annual anomaly (e_τ) which is explicitly neglected in SCA-V soil moisture retrievals. The temporal standard deviation of e_τ (σ_τ) expresses the statistical magnitude of this anomaly. Likewise, the serial lag-1 auto-covariance of e_τ (L_τ) captures the τ anomaly temporal auto-correlation strength. To be consistent with the TC soil moisture analysis, only days where sufficient soil moisture data were available for a TC analysis were considered when analyzing the magnitude and the temporal auto-correlation strength of τ .

3 Results

3.1 The total error of SCA-V and DCA retrievals

Figures 1 a and b plot the standard deviation of the total errors in SCA-V and DCA retrievals derived using the TC approach described above. Grey areas in Figure 1 (and Figure 2) represent pixels where TC-based SNR estimates of either SMAP or ASCAT retrievals are not significantly larger than 0.1 [-] (at $p = 0.05$ confidence; see section 2.1). Within CONUS, SCA-V and DCA retrievals exhibit similar spatial error distributions. In particular, surface soil moisture retrievals from both algorithms have the largest errors along both CONUS coasts and in heavily-vegetated portions of eastern CONUS. Conversely, lower errors are found in lightly-vegetated areas of west-central CONUS. As expected, spatial variations in total error variances for both algorithms are correlated with the spatial distribution of VWC (Figure 1 c). On average, SCA-V retrievals exhibit lower errors ($\sigma_{sca} = 0.017 \text{ m}^3/\text{m}^3$) than DCA retrievals ($\sigma_{dca} = 0.019 \text{ m}^3/\text{m}^3$). As noted above, these given values represent the standard deviation of serial errors in the scaled soil moisture anomalies. It does not include soil moisture seasonality error or multiplicative bias (scale) and is therefore not directly comparable to the unbiased root-mean-square error reported in *Chan et al.* [2016].

Figure 1 d plots differences in TC-derived total error (D_σ) between the two algorithms. Although the SCA-V soil moisture algorithm generally outperforms the DCA algorithm, SCA-V retrievals exhibit larger errors for areas of central CONUS (see red shading for areas of $D_\sigma > 0$ in Figure 1 d). Areas of superior DCA performance are also generally collocated with large deviations of τ from its climatological expectation - see the standard deviation of inter-annual τ anomalies (σ_τ) plotted in Figure 1 e. Recall that the primary difference between SCA-V and DCA retrievals lies in their treatment of τ . Therefore, by taking the differences between SCA-V and DCA errors, Figure 1 implies a direct link between inter-annual vegetation variability in τ and the localized degradation of SCA-V retrievals relative to DCA retrievals. Figure 1 f further confirms the relationship between D_σ and σ_τ . Here, D_σ is binned according to different percentiles of σ_τ , and a positive relationship between σ_τ and D_σ is found. A paired t-test was conducted for each bin to test if σ_{sca} and σ_{dca} are statistically different (at $p = 0.05$ confidence). Results of this test confirm that SCA-V significantly outperforms DCA when σ_τ is smaller than its 50th

CONUS percentile. Conversely, DCA significantly outperforms SCA-V when σ_τ is greater than its 75th CONUS percentile.

3.2 The autocorrelated errors of SCA-V and DCA retrievals

In addition to total error, we are also interested in examining the temporal auto-correlation of SMAP soil moisture retrieval errors using the TC analysis introduced in section 2.1. Overall, SCA-V retrievals show strong temporal error auto-covariance (L_{sca}) over high VWC regions in the eastern part of CONUS (Figure 2 a). In contrast, DCA generally has negligible auto-correlated error over the entire CONUS region (L_{dca} , Figure 2 b). On average, the serial lag-1 temporal auto-correlation (i.e., L_{sca}/σ_{sca}^2 and L_{dca}/σ_{dca}^2) is 0.364 [-] for SCA-V retrievals and 0.198 [-] for DCA retrievals. This indicates a larger amount of error auto-correlation in SCA-V retrievals relative to analogous DCA retrievals.

Similar to plotted values of σ_τ in Figure 1, the serial lag-1 auto-covariance of e_τ (L_τ) is plotted in Figure 2 d. As noted above, L_τ captures the degree of auto-correlation in inter-annual τ anomalies. Regions with strong positive D_L (i.e., large $L_{sca} - L_{dca}$ in Figure 2 c) tend to correspond to areas with relatively large L_τ values (Figure 2 d). In order to directly assess this tendency, Figure 2 e bins D_L according to different percentiles of L_τ , and demonstrates that D_L tends to increase with increased L_τ (Figure 2 e). In particular, when L_τ is larger than its 25th percentile, SCA-V retrievals have significantly ($p = 0.05$ confidence) higher error auto-covariances than comparable DCA retrievals. Taken as a whole, Figure 2 suggests that the neglect of inter-annual variability in τ by the SCA-V algorithm is associated with the presence of auto-correlated error in SCA-V soil moisture retrieval errors. It is also worth noting that areas where DCA outperforms SCA-V (in terms of total error, Figure 1 d) typically correspond to regions where SCA-V demonstrates highly auto-correlated errors (see e.g., the eastern part of CONUS in Figure 2 c). Therefore, the introduction of auto-correlated error in SCA-V retrievals via the neglect of inter-annual VWC variability appears to be a significant contributor to the uncertainty of SCA-V retrievals over these regions.

4 Discussion and conclusion

As our capacity to produce satellite-derived soil moisture products develops, it will become increasingly important to understand the statistical nature of errors present in

277 them. Here, we develop and apply a triple collocation (TC) technique to retrieve both the
278 total error variance and temporal auto-correlation coefficient for serial errors in SMAP soil
279 moisture retrievals derived from different retrieval algorithms. The current SMAP base-
280 line soil moisture retrieval algorithm (SCA) assumes accurate estimates of optical depth
281 (τ) can be obtained from a VWC seasonal climatology. Its potential alternative, the dual
282 channel algorithm or DCA, relaxes this assumption by simultaneously retrieving both soil
283 moisture and τ but is potentially less accurate with regards to the total retrieval error vari-
284 ance. Our central aim here is to better understand the temporal error characteristics of
285 retrievals acquired from both algorithms using a TC-based approach.

286 As expected, spatial patterns of error variance for both algorithms are closely related
287 to (annually-averaged) mean VWC conditions. Nevertheless, in terms of total error, SCA-
288 V slightly outperforms DCA, except for regions with large τ inter-annual variability (Fig-
289 ure 1). In addition, error in SCA-V retrievals contains significant auto-covariance within
290 the eastern CONUS. In contrast, errors in DCA retrievals exhibit generally negligible er-
291 ror auto-covariance. We demonstrate that SCA-V and DCA retrieval error auto-covariance
292 differences are closely related to the error auto-covariance of the inter-annual τ anomalies
293 neglected by SCA-V (Figure 2).

294 This leads us to the overall conclusion that both the presence of auto-correlated
295 errors in SCA-V retrievals and the superior performance of the DCA algorithm in cer-
296 tain areas can be linked directly to the neglect of inter-annual τ anomalies by the SCA-
297 V algorithm. Re-processing SCA-V retrievals using real-time (e.g., the 8-day or 16-day)
298 NDVI composites (to produce a real-time τ product) will likely significantly reduce auto-
299 correlated errors. However, these benefits will have to be weighed against the impact of
300 increased random errors in near-real-time NDVI composites required by SCA-V.

301 Although DCA contains larger total error than SCA-V, dual-polarization based soil
302 moisture retrievals should theoretically outperform single-polarization retrievals of soil
303 moisture. This departure from theoretical expectations is likely due to structural errors in
304 the simplified radiative transfer equation currently used for both DCA and SCA-V (e.g.,
305 the neglect of polarization dependence in b - see Entekhabi et al. 2014) which are not ex-
306 plicitly considered here. Future work will explore the possibility of applying TC-based
307 evaluation approaches to diagnose a wider range of structural problems affecting both re-
308 trieval approaches.

Acknowledgments

The authors thank Dr. Fan Chen for preparing the ASCAT data. This paper was partially supported by the NASA SMAP mission through award NNH12ZDA001N-SMAP. The SMAP data used in this study is available at <https://nsidc.org/data/smap/smap-data.html>, and the NLDAS2 data is available at <https://ldas.gsfc.nasa.gov/nldas/>.

References

- Albergel, C., C. Rüdiger, T. Pellarin, J.-C. Calvet, N. Fritz, F. Froissard, D. Suquia, A. Petitpa, B. Piguet, and E. Martin (2008), From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, *Hydrology and Earth System Sciences Discussions*, 12, 1323–1337, doi:10.5194/hess-12-1323-2008.
- Brown, M. E., V. Escobar, S. Moran, D. Entekhabi, P. E. O'Neill, E. G. Njoku, B. Doorn, and J. K. Entin (2013), NASA's soil moisture active passive (SMAP) mission and opportunities for applications users, *Bulletin of the American Meteorological Society*, 94(8), 1125–1128, doi:<https://doi.org/10.1175/BAMS-D-11-00049.1>.
- Chan, S., B. Rajat, H. Raymond, J. Tom, and K. John (2013), Ancillary data report: Vegetation water content, *JPL D-53061*.
- Chan, S., R. Bindlish, P. O'Neill, T. Jackson, E. Njoku, S. Dunbar, J. Chaubell, J. Piepmeier, S. Yueh, D. Entekhabi, et al. (2017), Development and assessment of the SMAP enhanced passive soil moisture product, *Remote Sensing of Environment*, doi:10.1016/j.rse.2017.08.025.
- Chan, S. K., R. Bindlish, P. E. O'Neill, E. Njoku, T. Jackson, A. Colliander, F. Chen, M. Burgin, S. Dunbar, J. Piepmeier, et al. (2016), Assessment of the SMAP passive soil moisture product, *IEEE Transactions on Geoscience and Remote Sensing*, 54(8), 4994–5007, doi:10.1109/TGRS.2016.2561938.
- Chen, F., W. T. Crow, A. Colliander, M. H. Cosh, T. J. Jackson, R. Bindlish, R. H. Reichle, S. K. Chan, D. D. Bosch, P. J. Starks, et al. (2017), Application of triple collocation in ground-based validation of soil moisture active/passive (SMAP) level 2 data products, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(2), 489–502, doi:10.1109/JSTARS.2016.2569998.
- Colliander, A., T. Jackson, R. Bindlish, S. Chan, N. Das, S. Kim, M. Cosh, R. Dunbar, L. Dang, L. Pashaian, et al. (2017), Validation of SMAP surface soil moisture prod-

ucts with core validation sites, *Remote Sensing of Environment*, 191, 215–231, doi:
<https://doi.org/10.1016/j.rse.2017.01.021>.

Crow, W., and M. Van den Berg (2010), An improved approach for estimating observa-
tion and model error parameters in soil moisture data assimilation, *Water Resources Re-
search*, 46(12), doi:10.1029/2010WR009402.

Crow, W. T., and M. T. Yilmaz (2014), The auto-tuned land data assimilation system (at-
las), *Water resources research*, 50(1), 371–385, doi:10.1002/2013WR014550.

Dorigo, W. A., K. Scipal, R. M. Parinussa, Y. Liu, W. Wagner, R. A. De Jeu, and
V. Naeimi (2010), Error characterisation of global active and passive microwave soil
moisture datasets, *Hydrology and Earth System Sciences*, 14(12), 2605–2616, doi:
10.5194/hess-14-2605-2010.

Draper, C., R. Reichle, R. de Jeu, V. Naeimi, R. Parinussa, and W. Wagner (2013),
Estimating root mean square errors in remotely sensed soil moisture over con-
tinental scale domains, *Remote Sensing of Environment*, 137, 288–298, doi:
<https://doi.org/10.1016/j.rse.2013.06.013>.

Entekhabi, D., E. G. Njoku, P. E. O’Neill, K. H. Kellogg, W. T. Crow, W. N. Edelstein,
J. K. Entin, S. D. Goodman, T. J. Jackson, J. Johnson, et al. (2010), The soil mois-
ture active passive (SMAP) mission, *Proceedings of the IEEE*, 98(5), 704–716, doi:
10.1109/JPROC.2010.2043918.

Entekhabi, D., S. Yueh, P. E. O’Neill, K. H. Kellogg, A. Allen, R. Bindlish, M. Brown,
S. Chan, A. Colliander, W. T. Crow, et al. (2014), SMAP handbook–soil moisture active
passive: Mapping soil moisture and freeze/thaw from space.

Gruber, A., C.-H. Su, S. Zwieback, W. Crow, W. Dorigo, and W. Wagner (2016), Recent
advances in (soil moisture) triple collocation analysis, *International Journal of Applied
Earth Observation and Geoinformation*, 45, 200–211, doi:10.1002/2013JD021043.

Koster, R. D., and M. J. Suarez (2001), Soil moisture memory in climate
models, *Journal of hydrometeorology*, 2(6), 558–570, doi:10.1175/1525-
7541(2001)002<0558:SMMICM>2.0.CO;2.

Koster, R. D., L. Brocca, W. T. Crow, M. S. Burgin, and G. J. De Lannoy (2016), Precip-
itation estimation using L-band and C-band soil moisture retrievals, *Water Resources
Research*, 52(9), 7213–7225, doi:10.1002/2016WR019024.

McColl, K. A., S. H. Alemohammad, R. Akbar, A. G. Konings, S. Yueh, and D. En-
tekhabi (2017), The global distribution and dynamics of surface soil moisture, *Nature*

374 *Geoscience*, 10(2), 100–104, doi:10.1038/ngeo2868.

375 Montzka, C., H. R. Boga, M. Zreda, A. Monerris, R. Morrison, S. Muddu, and
 376 H. Vereecken (2017), Validation of spaceborne and modelled surface soil moisture prod-
 377 ucts with cosmic-ray neutron probes, *Remote Sensing*, 9(2), 103, doi:10.3390/rs9020103.

378 Naeimi, V., K. Scipal, Z. Bartalis, S. Hasenauer, and W. Wagner (2009), An im-
 379 proved soil moisture retrieval algorithm for ERS and METOP scatterometer observa-
 380 tions, *IEEE Transactions on Geoscience and Remote Sensing*, 47(7), 1999–2013, doi:
 381 10.1109/TGRS.2008.2011617.

382 O'Neill, P., C. Steven, N. Eni, J. Tom, and B. Rajat (2015), Algorithm theoretical basis
 383 document level 2 & 3 soil moisture (passive) data products, *JPL D-66480*.

384 Ray, R. L., A. Fares, Y. He, and M. Temimi (2017), Evaluation and inter-comparison of
 385 satellite soil moisture products using in situ observations over texas, US, *Water*, 9(6),
 386 372, doi:10.3390/w9060372.

387 Scipal, K., T. Holmes, R. De Jeu, V. Naeimi, and W. Wagner (2008), A possible solution
 388 for the problem of estimating the error structure of global soil moisture data sets, *Geo-*
 389 *physical Research Letters*, 35(24), doi:10.1029/2008GL035599.

390 Wagner, W., G. Lemoine, and H. Rott (1999), A method for estimating soil moisture from
 391 ERS scatterometer and soil data, *Remote sensing of environment*, 70(2), 191–207, doi:
 392 10.1016/S0034-4257(99)00036-X.

393 Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. Wood, L. Luo, C. Alonge,
 394 H. Wei, J. Meng, et al. (2012), Continental-scale water and energy flux analysis and
 395 validation for the North American Land Data Assimilation System project phase 2
 396 (NLDAS-2): 1. intercomparison and application of model products, *Journal of Geophys-*
 397 *ical Research: Atmospheres*, 117(D3), doi:10.1029/2011JD016048.

398 Zwieback, S., W. Dorigo, and W. Wagner (2013), Estimation of the temporal
 399 autocorrelation structure by the collocation technique with an emphasis on
 400 soil moisture studies, *Hydrological sciences journal*, 58(8), 1729–1747, doi:
 401 <http://dx.doi.org/10.1080/02626667.2013.839876>.

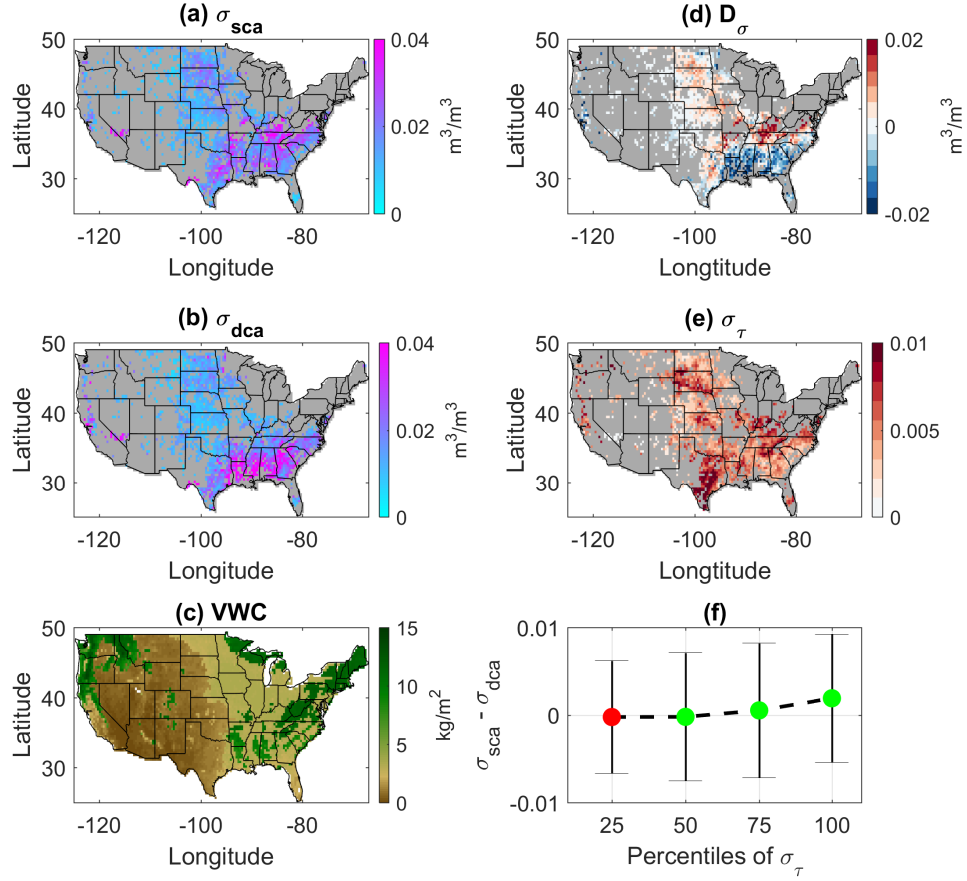


Figure 1. (a): The squared root of total error variance for SCA-V retrievals (a) and DCA retrievals (b); (c): average VWC conditions within the study period; (d): differences between square root of SCA-V and DCA total error variances ($D_{\sigma} = \sigma_{sca} - \sigma_{dca}$); (e): the standard deviation of inter-annual vegetation opacity (τ) anomalies (σ_{τ}); (f): D_{σ} as a function of σ_{τ} ; green (red) circles in (f) indicate that the mean of each D_{σ} bin is (not) significantly (at $p = 0.05$ confidence) different from zero; the error bar represents the inter-quartile range. Grey shading in maps indicates that the SNR of the SMAP or the ASCAT retrievals are not significantly (at $p = 0.05$ confidence) larger than 0.1 [-].

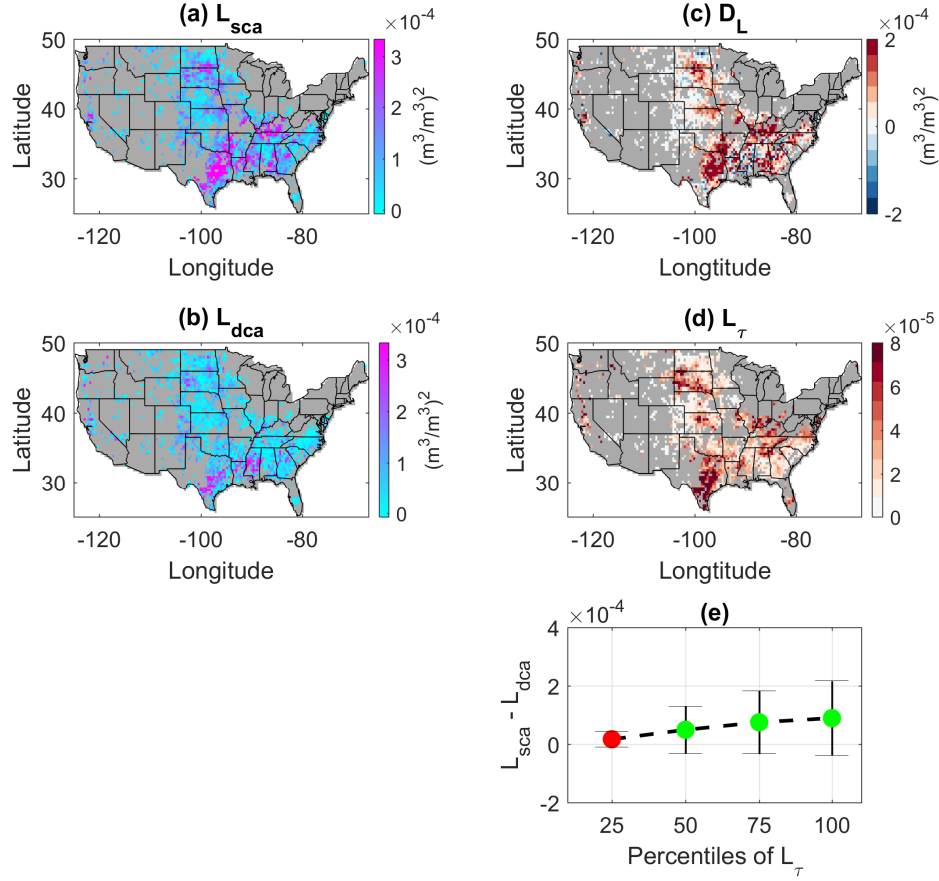


Figure 2. (a): Serial lag-1 error auto-covariance for SCA-V retrievals; (b): serial lag-1 error auto-covariance of DCA retrievals; (c): difference in the error auto-covariance of SCA-V and DCA retrievals ($D_L = L_{sca} - L_{dca}$); (d): serial lag-1 auto-covariance of inter-annual τ anomalies; (e): D_L as a function of L_τ ; green (red) circles in (e) indicate that the mean of each D_L bin are (not) significantly (at $p = 0.05$ confidence) different from zero; the error bar represents the inter-quartile range. Grey shading in maps indicates that the SNR of the SMAP or the ASCAT retrievals are not significantly ($p = 0.05$ confidence) larger than 0.1 [-].